National Animal Nutrition Program: Development of online feed composition tables

Andres Schlageter Tello
Multi-State Poultry Feeding and Nutrition Conference
May 21-23, 2019. Indianapolis, Indiana
https://animalnutrition.org/
Outline

- National Animal Nutrition Program (NANP)
- Feed composition tables
- Database
  - Literature dataset
  - Commercial laboratories dataset
- Work in progress
- Take home message

https://animalnutrition.org/
The National Animal Nutrition Program

https://animalnutrition.org/
Who Are We?

The National Animal Nutrition Program (NANP) serves as a forum to identify high-priority animal nutrition issues and provides an integrated and systemic approach to sharing, collecting, assembling, synthesizing, and disseminating science-based information, educational tools, and enabling technologies on animal nutrition that facilitate high-priority research among agricultural species.

Learn More
Coordinating Committee – M. Lindemann (UK)
• Oversee and coordinate the work of the feed composition and modeling groups, to advise the National Academies on critical national priorities, and to provide a forum to address research support needs

Feed Composition Committee – P. Miller (UNL)
• Bring together data and research resources on feed composition, to foster communication among those collecting feed composition information, and to facilitate efficiencies and consistencies in data collection and maintenance

Modeling Committee – M. Hanigan (VT)
• To serve the animal nutrition research community by improving the use of predictive technologies and tools, to best utilize available platforms, and to work with researchers to effectively share, combine, manage, manipulate, and analyze models and modeling information.

https://animalnutrition.org/
National Animal Nutrition Program
Feed Composition Committee

• Phil Miller (Chair), University of Nebraska (Swine)
• Andres Schlageter – University of Kentucky/University of Nebraska (Data)
• Ryan Dilger, University of Illinois (Poultry)
• Bill Dozier, Auburn University (Poultry)
• Mark Edwards, Cal Poly – San Luis Obispo (Equine)
• Alexander Hristov, Pennsylvania State University (Dairy)
• Brian Small, University of Idaho (Fish)
• Mark Nelson, Washington State University (Beef)
• Casey Bradley, DSM (Swine)
• William Weiss, The Ohio State University (Dairy)

https://animalnutrition.org/
National Animal Nutrition Program
Activities and “Products”

➢ Workshops and symposia:
  • Summit 2019: Producing Food with Animals. Sustainability, Efficiency and Security in US
  • ADSA 2019: Workshop, NANP nutrition models
  • ASAS 2019: Ruminant and Non-ruminant feed composition symposium

➢ Webpage resources:
  • NANP Publications (Abstracts, conferences papers and peer-reviewed papers)
  • Slides and videos from conferences
  • Codes

➢ Webpage databases:
  • Modeling database (performance of animals for modeling purposes)
  • Feed composition database

https://animalnutrition.org/
Feed Composition Tables

https://animalnutrition.org/
Feed composition tables

https://animalnutrition.org/

**Corn grain, dry** (Grain products)

- **Display Basis:** Dry Matter
- **DM Content (%):** 100

**Ingredients:**
- Corn grain, dry
- AAFCO: 48.4, Ground corn
- IFN: 4-02-961, Maize, grain ground
- EU: 1.2.1, maize

**Alternate Names:**
- Corn grain dry, ground, Corn grain, rolled, Corn, yellow dent
- Scientific Name: Zea mays sp. mays

<table>
<thead>
<tr>
<th>Nutrient (percentage of dry matter)</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>CV</th>
<th>10th percentile</th>
<th>90th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Dry Matter (DM, %)</td>
<td>400</td>
<td>88.10</td>
<td>2.00</td>
<td>2.27</td>
<td>85.61</td>
<td>90.30</td>
</tr>
<tr>
<td>Crude Protein (CP, %)</td>
<td>395</td>
<td>9.21</td>
<td>1.24</td>
<td>13.42</td>
<td>7.82</td>
<td>10.71</td>
</tr>
<tr>
<td>Crude Fiber (CF, %)</td>
<td>153</td>
<td>2.19</td>
<td>1.12</td>
<td>51.13</td>
<td>1.15</td>
<td>3.30</td>
</tr>
<tr>
<td>Ether Extract (EE, %)</td>
<td>306</td>
<td>3.99</td>
<td>1.07</td>
<td>26.72</td>
<td>3.01</td>
<td>5.00</td>
</tr>
<tr>
<td>Acid Ether Extract (AEE, %)</td>
<td>27</td>
<td>4.41</td>
<td>1.25</td>
<td>28.37</td>
<td>3.06</td>
<td>6.80</td>
</tr>
<tr>
<td>Ash (%)</td>
<td>227</td>
<td>1.60</td>
<td>1.06</td>
<td>66.24</td>
<td>1.16</td>
<td>1.90</td>
</tr>
<tr>
<td>Gross Energy (GE, kcal/kg)</td>
<td>90</td>
<td>4,543.60</td>
<td>223.19</td>
<td>4.91</td>
<td>4,396.97</td>
<td>4,636.66</td>
</tr>
</tbody>
</table>
Database and datasets

https://animalnutrition.org/
Current Database

Feed composition (127 nutrients)
- 7 Main constituents
- 28 Carbohydrates
- 22 Protein related nutrients
- 30 Fat related nutrients
- 23 Minerals
- 17 Vitamins

Feed names:
- NANP
- AAFCO
- International Feed Nomenclature
- European Union
- Definition

Nutrient definition (in progress)
- 127 Nutrients
- 128 nutritive values

Nutritive values (in progress)
- AME and Aa digestibility (Poultry)
- NE and Digestibility CP, Aa, P (Swine)
- NEI, NEm and degradability CP, NDF (Dairy)
- NEg, NEm and degradability CP (Beef)

https://animalnutrition.org/
Current Database: dataset

- Initially consolidating datasets from different NASEM/NRC committees
  - Literature datasets (swine, poultry)
  - Commercial laboratory datasets (beef, dairy)

- Current dataset (from literature only)
  - 4,807 feed samples
  - 99 different ingredients
  - 67 unique nutrients

- Near future dataset
  - ~2 million feed samples
  - 371 different ingredients
  - 137 unique nutrients

- Complete dataset is available under request

https://animalnutrition.org/
Dataset: Literature data

- Literature data are collected with literature reviews

- **Swine NRC 2012**
  - Articles published between 1998 and 2011
  - 2,777 feed samples
  - 147 unique ingredients
  - 67 nutrients

- **Database update (Poultry NASEM 2020?)**
  - Systematic literature review
  - Articles (> 30K) published between 2011 and 2018
  - 2,130 feed samples
  - 131 different ingredients
  - 91 unique nutrients

https://animalnutrition.org/
Dataset: Literature data

**Advantages**

- Better method to obtain unbiased data?
- Datasets are small and easy to manage (excel)
- Obtain information of nutrient not commonly analyzed (Amino acids, fatty acids, non-starch polysaccharides)

**Disadvantages**

- Values may not be representative of a feed (specially in feeds with low number of samples)
- **Time consuming (gather and review data)**
- Mistakes
  - Typing
  - Units % DM or % CP - % or g/Kg

https://animalnutrition.org/
Dataset: Commercial lab data

- **Beef NASEM 2016**
  - 3 commercial laboratories
  - 1.1 million feed samples
  - > 200 unique ingredients
  - 33 different nutrients

- **Dairy NASEM 2020?**
  - 4 commercial laboratories
  - 2.7 million feed samples
  - > 200 unique ingredients
  - 37 different nutrients

https://animalnutrition.org/
Dataset: Commercial lab data

- **Advantages**
  - Large datasets improves ensures analytes are more representative of the nutrient composition of feedstuffs being evaluated

- **Disadvantages**
  - Datasets have millions of data and have mistakes
  - File formats, data structure, and feed classifications differ among the feed testing laboratories.
  - Data management requires computer codes and high processing power.

Data management is difficult!!!

https://animalnutrition.org/
Commercial lab data: Understand your data

Histories for different nutrients in a feed initially identified as wheat grain

- **Dry matter %** = 89.7 ± 3.3
- **Crude protein %** = 14.6 ± 2.6
- **NDF %** = 12.4 ± 4.1

Wheat grain composition according NANP

- **Dry Matter %** = 89.5 ± 2.5
- **Crude Protein %** = 14.7 ± 2.3
- **NDF %** = 15.2 ± 8.0

https://animalnutrition.org/
Commercial lab data: Understand your data

Histograms for different nutrients in a feed initially identified as corn gluten meal

Dry matter % = 72.8 ± 23.1

Crude protein % = 43.2 ± 23.2

NDF % = 23.2 ± 16.3

Corn gluten meal composition according NANP

Dry Matter % = 91.5 ± 2.0

Crude Protein % = 63.2 ± 7.8

NDF % = 9.1 ± 6.6

https://animalnutrition.org/
Screening procedure for large dataset

Pre-screening

- Univariate
- PCA
- Clustering
- Cluster evaluation
- Data summary

https://animalnutrition.org/
Commercial lab data: Pre-screening

- **Delete not valid samples:**
  - Unidentified samples
  - Samples without values
  - Repeated samples
  - Samples referring to total mixed ration, concentrate, commercial brands or minerals
  - Non-feed samples (water and manure)

- **Standardize different sources:**
  - Standardize dataset structures (arrange columns in same order)
  - Standardize feed names
  - Standardize nutrient names and units

Pre-screening took 60% of time of the screening procedure

https://animalnutrition.org/
Pre-screening: Standardize different sources

Nutrient analytics in datasets provided by four commercial laboratories

<table>
<thead>
<tr>
<th>Laboratory 1</th>
<th>Laboratory 2</th>
<th>Laboratory 3</th>
<th>Laboratory 4</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM</td>
<td>DM</td>
<td>DM</td>
<td>DM</td>
<td>DM</td>
</tr>
<tr>
<td>Ash</td>
<td>Ash</td>
<td>Ash</td>
<td>Ash</td>
<td>Ash</td>
</tr>
<tr>
<td>Starch</td>
<td>Starch</td>
<td>Starch</td>
<td>Starch</td>
<td>Starch</td>
</tr>
<tr>
<td>Fat</td>
<td>Fat</td>
<td>Fat</td>
<td>Fat</td>
<td>Fat</td>
</tr>
<tr>
<td><strong>Total Fatty Acids</strong></td>
<td><strong>Total Fatty Acids</strong></td>
<td><strong>Total Fatty Acids</strong></td>
<td>Total Fatty Acids</td>
<td>Total Fatty Acids</td>
</tr>
<tr>
<td>NDF</td>
<td>NDF</td>
<td>NDF</td>
<td>NDF</td>
<td>NDF</td>
</tr>
<tr>
<td>ADF</td>
<td>ADF</td>
<td>ADF</td>
<td>ADF</td>
<td>ADF</td>
</tr>
<tr>
<td><strong>Water sol carb</strong></td>
<td><strong>Sugar</strong></td>
<td><strong>Water sol carb</strong></td>
<td><strong>Water sol carb</strong></td>
<td><strong>Water sol carb</strong></td>
</tr>
<tr>
<td>Ethanol Sol Carb</td>
<td>Ethanol Sol Carb</td>
<td>Ethanol Sol Carb</td>
<td>Ethanol Sol Carb</td>
<td>Ethanol Sol Carb</td>
</tr>
<tr>
<td>Lignin</td>
<td>Lignin</td>
<td>Lignin</td>
<td>Lignin</td>
<td>Lignin</td>
</tr>
<tr>
<td>Crude protein</td>
<td>Crude protein</td>
<td>Crude protein</td>
<td>Crude protein</td>
<td>Crude protein</td>
</tr>
<tr>
<td>Soluble Protein</td>
<td>Soluble Protein</td>
<td>Soluble Protein</td>
<td>Soluble Protein</td>
<td>Soluble Protein</td>
</tr>
<tr>
<td><strong>ADICP (% CP)</strong></td>
<td><strong>ADICP (% DM)</strong></td>
<td><strong>ADICP (% DM)</strong></td>
<td><strong>ADICP (% CP)</strong></td>
<td><strong>ADICP</strong></td>
</tr>
<tr>
<td><strong>NDICP (% CP)</strong></td>
<td><strong>NDICP (% DM)</strong></td>
<td><strong>NDICP (% DM)</strong></td>
<td><strong>NDICP (% CP)</strong></td>
<td><strong>NDICP</strong></td>
</tr>
<tr>
<td>NDFD24</td>
<td>NDFD30</td>
<td>NDFD30</td>
<td>sNDFD30</td>
<td>NDFD30</td>
</tr>
<tr>
<td>NDFD48</td>
<td>NDFD48</td>
<td>NDFD48</td>
<td>sNDFD48</td>
<td>NDFD48</td>
</tr>
<tr>
<td>NDFD240</td>
<td></td>
<td></td>
<td></td>
<td>NDFD240</td>
</tr>
<tr>
<td>Starch digest</td>
<td>Starch digest</td>
<td>Starch digest</td>
<td>Starch digest</td>
<td>Starch digest</td>
</tr>
</tbody>
</table>
Screening procedure for large dataset

1. Pre-screening
2. Univariate
3. PCA
4. Clustering
5. Cluster evaluation
6. Data summary

https://animalnutrition.org/
Univariate: Delete outliers

- 3.5σ
- 3σ - 2σ - 1σ μ 1σ 2σ 3σ 3.5σ

68% of data
95% of data
99.7% of data

Deleted
Deleted

https://animalnutrition.org/
Principal Component Analysis (PCA)

• Reduce dimensionality (In this case dimensions are different nutrients used to classify feeds, i.e. DM, CP, NDF etc)
• Retain as much variation as possible
• Linear transformation of the original variables
• PCA + Clustering is a widely used protocol, provide better clusters when compared with using

https://animalnutrition.org/
Hierarchical clustering
Grouping similar data into a single group

https://animalnutrition.org/
Hierarchical clustering

https://animalnutrition.org/
Hierarchical clustering

https://animalnutrition.org/
Hierarchical clustering
Hierarchical clustering

https://animalnutrition.org/
Hierarchical clustering

https://animalnutrition.org/
Hierarchical clustering: When to stop?
Pseudo F and Pseudo t to select optimal number of clusters

https://animalnutrition.org/
Univariate: Delete outliers from clusters

- 3.5σ

68% of data

95% of data

99.7% of data

3.5σ

Deleted

Deleted

https://animalnutrition.org/
Screening procedure for large datasets

- Pre-screening
- Univariate
- PCA
- Clustering
- Cluster evaluation

Data summary

https://animalnutrition.org/
Cluster Evaluation

https://animalnutrition.org/
Cluster Evaluation

Cluster evaluation procedure was performed according authors’ expertise in feed composition.

Some of the procedures used in cluster evaluation included:

- Repeating the screening procedure (the automated statistical screening procedure was repeated between 1 and 5 times).
- Merging files initially identified as different feeds to create a new input file.
- Merging clusters generated from different initial input files.
- Retrieving clusters removed by the procedure.
- Manual manipulation of cluster’s datasets.

https://animalnutrition.org/
Clustering example

Initial dataset identified as Soybean meal

Output clusters identified as:
- Soybean meal solvent extracted
- Soybean meal mechanical extracted
Clustering example

Output clusters for a dataset initially identified as Bakery byproduct

- Bakery byproduct, cereal
- Bakery byproduct, cookies
- Bakery byproduct, bread waste

https://animalnutrition.org/
Screening procedure for large datasets

- Data gathering
- Pre-screening
- Univariate
- PCA
- Clustering
- Cluster evaluation
- Data summary

https://animalnutrition.org/
Data Summary

https://animalnutrition.org/
Screening procedure for large datasets

**Advantages**
- Eliminated outlier data points
- The procedure was helpful to correctly classify feeds:
  - Oilseeds and oilseed meals according to crude fat values
  - By-products and grains as dry or wet according dry matter (DM) content
  - Identify misclassified samples in feed similar names
  - Classify forages according to storage method
  - Classify forages according maturity stage
  - Classify samples as predominantly-grass or predominantly-legume

**Disadvantages**
- About 50% of data were deleted
- Time consuming

https://animalnutrition.org/
Work in progress
New methods to classify feeds
Decision tree

- Classification algorithm
- Split the main dataset into two or more homogeneous datasets based on most significant differentiator/splitter in input variables.
- Decision trees are easy to create and understand. However, they tend to over fit.

https://animalnutrition.org/
New methods to classify feeds
Random forest

- Classification algorithm.

- Ensemble method: combine several base models in order to produce a better predictive (classifier) model.

- Each decision tree classify a sample in a group. The sample is classified according the aggregated results of each decision tree.

https://animalnutrition.org/
New methods to classify feeds
Preliminary results

Precision of decision tree and random forest algorithms to classify different corn grain feeds

<table>
<thead>
<tr>
<th>Feed</th>
<th>Decision tree</th>
<th>Random forest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corn germ</td>
<td>89%</td>
<td>92%</td>
</tr>
<tr>
<td>Corn germ meal</td>
<td>96%</td>
<td>99%</td>
</tr>
<tr>
<td>Corn gluten feed, dry</td>
<td>99%</td>
<td>100%</td>
</tr>
<tr>
<td>Corn gluten feed, wet</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Corn gluten meal</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Corn grain, dry</td>
<td>90%</td>
<td>93%</td>
</tr>
<tr>
<td>Corn grain, high moisture</td>
<td>99%</td>
<td>99%</td>
</tr>
<tr>
<td>Corn grain steam flaked</td>
<td>72%</td>
<td>81%</td>
</tr>
<tr>
<td>Corn hominy feed</td>
<td>89%</td>
<td>94%</td>
</tr>
<tr>
<td>Corn screenings</td>
<td>72%</td>
<td>80%</td>
</tr>
</tbody>
</table>

[https://animalnutrition.org/](https://animalnutrition.org/)
Collecting information about:

- **Non-Starch polysaccharides**
  - Systematic literature review

- **Nutritive values for different species. For poultry**...
  - Apparent metabolizable energy and apparent metabolizable energy nitrogen corrected
  - Apparent ileal digestibility for amino acids
  - Standard ileal digestibility for amino acids

- **Include more species**
  - Fish
  - Horses
  - Small ruminants

https://animalnutrition.org/
The National Animal Nutrition Program (NANP) is creating multispecies feed composition datasets and tables.

Feed composition tables are constructed using nutrition information from literature and commercial laboratories.

All information created by NANP can be found at https://animalnutrition.org/

Datasets are available under request.